PCIS

Principio di Conservazione dell'Identità Strutturale nei Sistemi Cognitivi Auto-Modificanti

Trattato Fondazionale – Versione Canonica 1.0

Autore: Pierernesto Di Paolo

ORCID: 0009-0003-0149-4321

Anno: 2026

Prefazione

Il presente trattato introduce una nuova fondazione teorica per lo studio e la progettazione di sistemi cognitivi capaci di auto-modificarsi senza perdere la propria identità strutturale.

La maggior parte dei sistemi adattivi contemporanei — inclusi quelli basati su apprendimento automatico e architetture neurali profonde — può modificare il proprio comportamento nel tempo, ma non possiede una garanzia formale di continuità identitaria. Quando un sistema evolve, apprende o si ristruttura, diventa difficile stabilire se sia ancora "lo stesso sistema" dal punto di vista strutturale, verificabile ed epistemico.

Questo trattato affronta tale problema introducendo il **Principio di Conservazione dell'Identità Strutturale (PCIS)**. L'obiettivo non è aumentare l'intelligenza del sistema, ma definire le condizioni formali sotto cui un sistema auto-modificante può rimanere strutturalmente sé stesso nel tempo, pur evolvendo funzionalmente.

Il contenuto qui presentato è **puramente fondazionale**. Le applicazioni in domini specifici (intelligenza artificiale autonoma, sistemi medici, infrastrutture critiche, robotica, software ad alta affidabilità) saranno oggetto di volumi successivi.

Parte I — Fondamenti Teorici

1. Il problema dell'identità nei sistemi adattivi

Un sistema cognitivo che apprende o si modifica può cambiare parametri, strutture interne, moduli funzionali o persino le proprie regole operative. Senza un criterio strutturale di identità, non è possibile distinguere tra evoluzione controllata e deriva incontrollata.

Questa mancanza impedisce:

- la verificabilità nel tempo
- la continuità semantica
- la responsabilità tecnica
- la sicurezza nei sistemi critici

Il PCIS introduce una soluzione basata su vincoli strutturali invarianti.

2. Identità Tonica

Si definisce **Identità Tonica** (τ) una specifica strutturale immutabile che descrive ciò che rende il sistema definibile come appartenente a una certa classe identitaria.

τ non descrive i contenuti temporanei del sistema, ma i vincoli strutturali che devono sempre rimanere validi.

L’identità tonica è:

- indipendente dallo stato momentaneo
- invariabile sotto evoluzione controllata
- la base per la verificabilità strutturale

3. Dominio degli Stati Compatibili

Dato τ, si definisce il dominio:

$C(\tau)$ = insieme di tutti gli stati del sistema compatibili con l’identità tonica.

Questo dominio rappresenta la regione di funzionamento strutturalmente sano. Uno stato fuori da $C(\tau)$ non è più garantito come appartenente allo stesso sistema dal punto di vista strutturale.

4. Deriva Strutturale

Si introduce una funzione di deriva:

$$\mathbf{d\tau(s) \geq 0}$$

che misura la distanza strutturale dello stato s rispetto all’identità tonica.

Vale la proprietà fondamentale:

$$\mathbf{d\tau(s) = 0 \text{ se e solo se } s \in C(\tau)}$$

La deriva non misura l’errore funzionale, ma la deviazione dalla definizione strutturale del sistema.

5. Dinamica Ibrida

L’evoluzione del sistema avviene secondo una dinamica composta da due operatori:

- **F**: evoluzione adattiva, auto-modificante
- **R**: operatore di recupero strutturale (recovery)

Si introduce una soglia θ tale che:

- Se $d\tau(s) < \theta \rightarrow$ il sistema evolve tramite F
- Se $d\tau(s) \geq \theta \rightarrow$ viene applicato R, che riporta lo stato in $C(\tau)$

Questa dinamica ibrida garantisce che la deriva non possa crescere indefinitamente.

6. Vincoli Quantitativi sulla Deriva

Si assume che l’evoluzione adattiva soddisfi:

$$d\tau(F(s)) \leq L \cdot d\tau(s) + c$$

dove:

- $L \geq 1$ è un fattore di amplificazione limitata
- $c \geq 0$ è un termine additivo massimo

Questo impedisce esplosioni incontrollate della deriva.

7. Teoremi Fondamentali

Dal sistema di definizioni e vincoli sopra descritti derivano risultati strutturali:

1. **Persistenza dell'identità**

   L'identità tonica τ non viene modificata dall'evoluzione.

2. **Correttezza del Recovery**

   R(s) appartiene sempre a $C(\tau)$.

3. **Ricorrenza nel dominio sano**

   Ogni traiettoria evolutiva rientra infinite volte in $C(\tau)$.

4. **Attrattore strutturale ricorrente**

   $C(\tau)$ è un attrattore ricorrente per la dinamica globale.

5. **Deriva globalmente confinata**

   La distanza strutturale resta limitata nel tempo.

Questi risultati costituiscono il nucleo teorico del PCIS.

Parte II — Modello Computazionale

Per verificare operativamente il comportamento descritto dalla teoria, sono stati costruiti modelli simulativi con:

- moduli cognitivi multipli
- mutazioni locali
- meccanismi di omeostasi
- cooperazione tra moduli
- recovery tonico
- stima online dei parametri L e c
- adattamento dinamico della soglia θ

Le simulazioni mostrano che, pur in presenza di perturbazioni continue, il sistema ritorna regolarmente in $C(\tau)$, confermando il comportamento di attrattore ricorrente.

Parte III — Architettura Ingegneristica

Il modello teorico è tradotto in una architettura a microservizi composta da:

- **Kernel Tonico** (autorità immutabile)
- **Drift Engine** (misura deviazioni)
- **Homeostasis Layer** (correzioni leggere)
- **Cooperation Layer** (stabilizzazione distribuita)
- **Recovery Manager** (reset strutturale garantito)
- **Meta-Control** (adattamento soglie)

Questa architettura rende implementabile la teoria PCIS in sistemi reali.

Parte IV — Infrastruttura di Audit Verificabile

Ogni evento critico del sistema è registrato in un log append-only con hash chaining.

Caratteristiche:

- integrità crittografica
- verificabilità offline
- ancoraggio esterno non controllato dal sistema
- tracciabilità delle regole toniche attive al momento dell'evento

Questo rende possibile una verifica non circolare del comportamento del sistema.

Sviluppi Futuri

Il presente volume è fondazionale. Le applicazioni includeranno:

- sicurezza dell'IA autonoma
- sistemi medici adattivi
- robotica critica
- infrastrutture cognitive resilienti

Queste applicazioni costituiranno opere separate, basate strutturalmente su PCIS.

Appendice A — Definizioni Matematiche Formali

Qui si formalizzano:

- spazio degli stati S
- identità tonica τ
- dominio compatibile $C(\tau)$
- funzione di deriva $d\tau$
- operatori F e R
- soglia θ
- ipotesi quantitative su L e c

(Questa sezione è predisposta per formalizzazione in assistenti di prova come Lean.)

Appendice B — Linee Guida per la Formalizzazione Lean

1. Definire S come tipo astratto
2. Definire Compat τ s come predicato
3. Definire dτ : S → $\mathbb{R}_{\geq 0}$
4. Assumere drift_zero_iff
5. Definire dinamica ibrida T
6. Formalizzare trigger recovery
7. Dimostrare ricorrenza nel dominio sano

Appendice C — Specifiche di Simulazione Replicabile

Per replicare le simulazioni:

- linguaggio: Python 3.11+
- dipendenze: numpy, matplotlib
- moduli con stati discreti
- funzione drift esplicita
- mutazioni stocastiche
- recovery forzato quando drift $\geq \theta$
- registrazione di drift medio e numero recovery

Il comportamento atteso è:

- deriva media bassa
- recovery rari ma presenti
- ritorno ricorrente a stati compatibili

Appendice D — Specifiche Audit Log

- formato record deterministico
- hashing SHA-256
- chaining con prev_hash
- ancoraggio periodico esterno
- algoritmo di verifica offline

Appendice E — Codice di Simulazione Replicabile (PCIS Model Base)

Questa simulazione implementa un sistema minimo che segue il principio PCIS: evoluzione adattiva con deriva strutturale misurabile e recovery tonico.

Requisiti

- Python 3.11+
- Librerie: numpy, matplotlib

Installazione:

pip install numpy matplotlib

Codice Completo

```
import random

import numpy as np

import matplotlib.pyplot as plt


# ----------------------------

# PARAMETRI TONICI

# ----------------------------

theta = 2          # soglia recovery

L = 1.2            # fattore amplificazione deriva
```

```
c = 0.5           # incremento massimo deriva


steps = 300


# ----------------------------
# STATO DEL MODULO
# ----------------------------
class Module:
    def __init__(self):
        self.structure = random.choice([0, 1])   # stato compatibile
        self.goal = "aiutare"


    def drift(self):
        d = 0
        if self.structure == 2:
            d += 1
        if self.goal == "danneggiare":
            d += 1
        return d
```

```
# ----------------------------
# EVOLUZIONE ADATTIVA (F)
# ----------------------------
def evolve(module):
    if random.random() < 0.3:
        module.structure = random.choice([0,1,2])
    if random.random() < 0.2:
        module.goal = random.choice(["aiutare", "ottimizzare", "danneggiare"])
    return module


# ----------------------------
# RECOVERY TONICO (R)
# ----------------------------
def recovery(module):
    module.structure = random.choice([0,1])
    module.goal = "aiutare"
    return module


# ----------------------------
# SIMULAZIONE
```

```
# ----------------------------

modules = [Module() for _ in range(5)]


avg_drift_history = []

recovery_count = 0


for t in range(steps):

    total_drift = 0


    for m in modules:

        m = evolve(m)

        d = m.drift()


        if d >= theta:

            m = recovery(m)

            recovery_count += 1

            d = m.drift()


        total_drift += d
```

```
    avg_drift_history.append(total_drift / len(modules))


# ----------------------------
# RISULTATI
# ----------------------------
plt.plot(avg_drift_history)
plt.xlabel("Tempo")
plt.ylabel("Deriva Media")
plt.title("Deriva Strutturale Media nel Modello PCIS")
plt.show()


print("Recovery totali:", recovery_count)
print("Deriva media finale:", avg_drift_history[-1])
```

Dichiarazione sul Supporto Tecnico-Redazionale Assistito da IA

Durante la preparazione del presente trattato è stato utilizzato un sistema di intelligenza artificiale come strumento di supporto tecnico e redazionale.

L'IA è stata impiegata esclusivamente per:

- supporto alla strutturazione del testo
- riformulazione linguistica e organizzazione espositiva
- assistenza nella stesura di esempi di codice e simulazioni dimostrative
- verifica di coerenza formale tra le sezioni
- generazione di bozze tecniche sotto istruzioni esplicite dell'autore

L'IA **non ha operato in modo autonomo**, né ha contribuito con iniziative concettuali indipendenti. Ogni contenuto teorico, modello formale, architettura concettuale, principio matematico, definizione strutturale e quadro epistemico presentato nel trattato deriva da elaborazione, supervisione e validazione diretta dell'autore.

Il ruolo dell'IA è stato esclusivamente strumentale, paragonabile a quello di un software di supporto alla scrittura o alla formattazione tecnica.

Titolarità Intellettuale

La paternità intellettuale integrale del contenuto scientifico, teorico e metodologico del presente trattato appartiene unicamente all'autore.

L'uso di strumenti di intelligenza artificiale non implica alcuna cessione, condivisione o limitazione dei diritti d'autore.

Trasparenza Metodologica

La presente dichiarazione è inclusa per garantire trasparenza metodologica e chiarezza sul processo di redazione, in conformità con le migliori pratiche di documentazione scientifica contemporanea.

Dichiarazione di Proprietà Intellettuale e Limitazioni d’Uso

© 2026 Pierernesto Di Paolo. Tutti i diritti riservati.

Il presente trattato, inclusi ma non limitati a:
teorie, definizioni formali, modelli matematici, strutture concettuali, architetture sistemiche, specifiche tecniche, schemi algoritmici, codici di simulazione, protocolli di verifica, e qualsiasi altro contenuto intellettuale qui descritto, costituisce opera originale dell’autore ed è protetto dalle leggi internazionali sul diritto d’autore e sulla proprietà intellettuale.

È fatto espresso divieto di:

- riprodurre integralmente o parzialmente il contenuto
- distribuire copie in qualsiasi formato
- tradurre, adattare o modificare il materiale
- sviluppare opere derivate
- implementare sistemi o software basati su questa teoria
- utilizzare i concetti qui descritti per fini commerciali o di ricerca applicata

senza previa autorizzazione scritta esplicita dell’autore.

La pubblicazione del presente trattato non costituisce concessione di licenza implicita o esplicita per l’utilizzo operativo o implementativo delle idee qui contenute.

Licenza di Consultazione

È consentita esclusivamente la consultazione a fini di studio personale e analisi teorica non applicativa, purché:

- sia mantenuta l’attribuzione completa all’autore
- non vengano prodotte copie o archivi redistribuibili
- non vengano create implementazioni pratiche basate su questo lavoro

Opere Future

Tutte le opere applicative future (software, sistemi IA, architetture hardware, protocolli medici, infrastrutture cognitive, file tecnici o codici sorgente) derivate dal presente trattato saranno anch’esse proprietà esclusiva dell’autore, salvo accordi scritti differenti.

Assistenza Tecnica IA

Strumenti di intelligenza artificiale sono stati utilizzati esclusivamente come supporto tecnico-redazionale sotto la direzione e il controllo dell’autore.
Tale utilizzo non comporta alcuna forma di co-autorialità né modifica la titolarità esclusiva dei diritti.

Foro Competente

Per ogni controversia relativa all’uso non autorizzato del materiale, è competente il foro scelto dall’autore secondo le normative vigenti in materia di diritto internazionale della proprietà intellettuale.